

# LEGGE DEI GRANDI NUMERI

E. DI NARDO

## 1. LEGGE EMPIRICA DEL CASO E IL TEOREMA DI BERNOULLI

In diverse occasioni, abbiamo menzionato che la nozione intuitiva di probabilità si basa sulla seguente assunzione: se in  $n$  sperimentazioni (identiche) un certo evento  $A$  si verifica  $m \leq n$  volte, allora il rapporto  $m/n$  deve essere *vicino* a  $P(A) = p$  per  $n$  sufficientemente elevato. Questa assunzione viene espressa dalla cosiddetta *legge empirica del caso*: in una successione di prove fatte nelle stesse condizioni, la frequenza relativa dell'evento si avvicina alla probabilità dell'evento stesso e l'approssimazione tende a migliorare con l'aumentare del numero delle prove. La legge empirica del caso è alla base della cosiddetta impostazione frequentista del calcolo delle probabilità. Nel 1713, Bernoulli dimostrò un teorema che aveva molte "somiglianze" con la legge empirica del caso.

**Teorema 1.1.** *Sia  $\{X_n\}$  una successione di v.a. di Bernoulli i.i.d. con parametro  $p$ . Vale che*

$$(1.1) \quad \lim_{n \rightarrow \infty} P(|S_n - np| < z) = 0 \quad \forall z > 0$$

$$(1.2) \quad \lim_{n \rightarrow \infty} P\left(\left|\frac{S_n}{n} - p\right| < z\right) = 1 \quad \forall z > 0$$

dove  $S_n = X_1 + \dots + X_n$ .

Per capire il significato dei limiti in (1.1) e (1.2) diamo la definizione di convergenza in probabilità.

**Definizione 1.2.** Sia  $\{X_n\}$  una successione di v.a. definite sullo stesso spazio di probabilità. Diremo che  $X_n$  converge in probabilità a  $X$ , e scriveremo  $X_n \xrightarrow{P} X$ , se e solo se fissato un  $\varepsilon > 0$  e indicato con  $A_n = \{\omega \in \Omega : |X_n(\omega) - X(\omega)| < \varepsilon\}$  risulta

$$\lim_{n \rightarrow \infty} P(A_n) = 1.$$

In base a questa definizione il risultato (1.2) si riscrive come

$$\frac{S_n}{n} \xrightarrow{P} p.$$

I due risultati sembrano in contraddizione: il primo dice che pur di attendere un tempo sufficientemente lungo, la successione delle probabilità degli insiemi contenenti le realizzazioni del processo di Bernoulli per le quali il numero di successi al passo  $n$  si mantiene nell'intervallo  $(np - z, np + z)$  tende a zero; il secondo invece asserisce che la successione delle probabilità degli insiemi contenenti le realizzazioni del processo di Bernoulli per le quali la frequenza relativa di occorrenza del successo

si mantiene nella fascia  $(p - z, p + z)$  tende a 1, pur di attendere un tempo sufficientemente lungo. In genere i giocatori d'azzardo tengono a mente il secondo risultato piuttosto che il primo. Per un giocatore che punta sul successo, l'equazione (1.1) è decisamente una cattiva notizia: se  $S_n$  rappresenta il capitale guadagnato e  $np$  quello atteso, l'equazione dice che il primo non resta in un opportuno intorno del secondo. Anche la seconda equazione non è una buona notizia: per  $n$  sufficientemente largo, la frequenza relativa dell'evento tende a  $p$  ma non è detto che resti in questo intorno per un numero  $k$  di prove successive, inoltre c'è sicuramente una realizzazione del processo di Bernoulli per la quale la frequenza di occorrenza vale sempre 1. Vi sono varie dimostrazioni della (1.1): ad esempio si può utilizzare il teorema centrale del limite. Per dimostrare la (1.2), usiamo la disuguaglianza di Cebicev.

**Proposizione 1.3.** *Per  $r > 0$  e  $a > 0$  si ha*

$$P(|X| < a) \geq 1 - \frac{E[|X|^r]}{a^r}.$$

*Proof.*

$$E[|X|^r] = \int_{\mathbb{R}} |x|^r dF(x) \geq \int_{|x| \geq a} |x|^r dF(x) \geq a^r P(|X| \geq a).$$

□

Se esiste finito il momento secondo della v.a.  $X$  allora per  $t > 0$  si ha

$$P(|X - \mu| < t\sigma) \geq 1 - \frac{E[|X - \mu|^2]}{t^2\sigma^2} = 1 - \frac{1}{t^2}.$$

Nel caso della (1.2) si ha

$$P\left(\left|\frac{S_n}{n} - p\right| < z\right) \geq 1 - \frac{1}{z^2} E\left[\left(\frac{S_n}{n} - p\right)^2\right] = 1 - \frac{npq}{n^2 z^2}$$

da cui la convergenza a 1 per  $n$  che tende all'infinito.

La seconda parte del teorema di Bernoulli viene anche chiamata legge (debole) dei grandi numeri. Si legge in alcuni testi che la legge empirica del caso è anche nota col nome di legge dei grandi numeri. Questa affermazione non è corretta, perché in realtà si tratta di due enunciati concettualmente molto diversi. La legge dei grandi numeri è un teorema (cioè un'affermazione dimostrabile): a dire il vero, sono state formulate diverse leggi dei grandi numeri, che si collocano a diversi gradi di generalità. La legge empirica del caso è, invece, tutt'altro che un teorema: diciamo che è una verità rilevabile sperimentalmente, ma su questo si potrebbe in realtà discutere per giornate intere.

## 2. LEGGE DEBOLE DEI GRANDI NUMERI

La legge debole dei grandi numeri estende il teorema di Bernoulli a v.a. con distribuzione diversa da quella di Bernoulli.

**Teorema 2.1** (Teorema di Kincin). *Sia  $\{X_n\}$  una successione di v.a. i.i.d. con media finita. Allora*

$$\frac{S_n}{n} \xrightarrow{P} \mu$$

dove  $E[X_i] = \mu, \forall i$  e  $S_n = X_1 + X_2 + \dots + X_n$ .

Non avendo ipotesi sulla esistenza del momento secondo, non si può utilizzare la disuguaglianza di Cebicev per la dimostrazione del teorema.

*Proof.* Dimostriamo l'asserto nel caso  $\mu = 0$  poichè il caso  $\mu \neq 0$  si riconduce a questo considerando le v.a.  $Y_n = X_n - \mu$ . Sia  $\phi(t)$  la funzione caratteristica di una generica  $X_i$ . Risulta

$$\phi_{S_n/n}(t) = \left[ \phi\left(\frac{t}{n}\right) \right]^n.$$

Poiché  $X$  ha media nulla,  $\phi(t) = 1 + R_1(t)$ , dove  $R_1(t)$  è un infinitesimo di ordine superiore a  $t$  e quindi

$$\phi_{S_n/n}(t) = \left[ 1 + R_1\left(\frac{t}{n}\right) \right]^n = \left[ 1 + \frac{1}{n} nR_1\left(\frac{t}{n}\right) \right]^n.$$

Poiché

$$\lim_{n \rightarrow \infty} nR_1\left(\frac{t}{n}\right) = 0$$

segue che

$$\lim_{n \rightarrow \infty} \phi_{S_n/n}(t) = 1.$$

La funzione caratteristica pari a 1 corrisponde alla v.a. degenera nulla  $P(X = 0) = 1$ , sicché abbiamo provato che

$$\frac{S_n}{n} \xrightarrow{d} 0.$$

La v.a. degenera ha funzione di ripartizione data da

$$F(x) = \begin{cases} 0 & x < 0 \\ 1 & x \geq 0 \end{cases}$$

di conseguenza per  $z > 0$  si ha

$$P\left(\left|\frac{S_n}{n}\right| < z\right) = P\left(-z < \frac{S_n}{n} < z\right) \geq P\left(-z < \frac{S_n}{n} \leq \frac{z}{2}\right) = F_n\left(\frac{z}{2}\right) - F_n(-z)$$

il cui limite tende a 1. □

Questo risultato può essere rafforzato, nel senso che si può dare l'asserto del teorema utilizzando una convergenza più restrittiva di quella in probabilità.

### 3. LEGGE FORTE DEI GRANDI NUMERI

**Definizione 3.1.** Sia  $\{X_n\}$  una successione di v.a. definite sullo stesso spazio di probabilità. Diremo che  $X_n$  converge quasi certamente (q.c.) a  $X$ , e scriveremo  $X_n \xrightarrow{q.c.} X$ , se e solo se, indicato con  $A = \{\omega \in \Omega : \lim_{n \rightarrow \infty} X_n(\omega) = X(\omega)\}$  si ha  $P(A) = 1$ .

**Teorema 3.2** (Teorema di Kolmogorov). *Sia  $\{X_n\}$  una successione di v.a. i.i.d. con media finita. Allora*

$$\frac{S_n}{n} \xrightarrow{q.c.} \mu$$

dove  $E[X_i] = \mu, \forall i$  e  $S_n = X_1 + X_2 + \dots + X_n$ .

La legge forte dei grandi numeri asserisce che il rapporto  $S_n/n-p$  diventa e resta piccolo, per  $n$  grande. Inoltre si può dimostrare che l'esistenza della media finita è condizione anche necessaria per la convergenza. Non daremo la dimostrazione del teorema, che è piuttosto complicata. Invece faremo vedere perché questa versione del teorema è più forte di quella del paragrafo precedente.

Premettiamo il seguente risultato.

**Teorema 3.3.**

$$X_n \xrightarrow{q.c.} X \Leftrightarrow \forall \epsilon > 0 \lim_{n \rightarrow \infty} P(\cap_{m=n}^{\infty} \{\omega \in \Omega : |X_m(\omega) - X(\omega)| < \epsilon\}) = 1.$$

*Proof.* Si tratta di provare che indicato con  $A$  l'evento

$$A = \{\omega \in \Omega : \lim_{n \rightarrow \infty} X_n(\omega) = X(\omega)\}$$

risulta

$$(3.1) \quad A = \bigcap_{k=1}^{\infty} \bigcup_{n=1}^{\infty} \bigcap_{m=n}^{\infty} \left\{ \omega \in \Omega : |X_m(\omega) - X(\omega)| < \frac{1}{k} \right\}.$$

Infatti

$$\begin{aligned} A &= \{\omega \in \Omega : \forall \epsilon > 0, \exists n, \forall m \geq n |X_m(\omega) - X(\omega)| < \epsilon\} \\ &= \left\{ \omega \in \Omega : \forall k > 0, \exists n, \forall m \geq n |X_m(\omega) - X(\omega)| < \frac{1}{k} \right\} \end{aligned}$$

con  $k$  intero. È evidente che

$$\{\omega \in \Omega : \forall k > 0 \dots\} = \bigcap_{k=1}^{\infty} \{\omega \in \Omega : \dots\}$$

$$\{\omega \in \Omega : \exists n > 0 \dots\} = \{\omega \in \Omega : \dots\}$$

da cui segue la (3.1). Osserviamo che gli eventi

$$A_k = \bigcup_{n=1}^{\infty} \bigcap_{m=n}^{\infty} \left\{ \omega \in \Omega : |X_m(\omega) - X(\omega)| < \frac{1}{k} \right\}$$

sono decrescenti e quindi  $\cap_{k=1}^{\infty} A_k = \lim_{k \rightarrow \infty} A_k$  mentre gli eventi

$$B_n = \bigcap_{m=n}^{\infty} \left\{ \omega \in \Omega : |X_m(\omega) - X(\omega)| < \frac{1}{k} \right\}$$

sono crescenti, ragion per cui  $\cup_{n=1}^{\infty} B_n = \lim_{n \rightarrow \infty} B_n$ . Per la continuità della probabilità,

$$\begin{aligned} &P\left(\bigcap_{k=1}^{\infty} \bigcup_{n=1}^{\infty} \bigcap_{m=n}^{\infty} \left\{ \omega \in \Omega : |X_m(\omega) - X(\omega)| < \frac{1}{k} \right\}\right) \\ &= P\left(\lim_{k \rightarrow \infty} \lim_{n \rightarrow \infty} \bigcap_{m=n}^{\infty} \left\{ \omega \in \Omega : |X_m(\omega) - X(\omega)| < \frac{1}{k} \right\}\right) \\ &= \lim_{k \rightarrow \infty} \lim_{n \rightarrow \infty} P\left(\bigcap_{m=n}^{\infty} \left\{ \omega \in \Omega : |X_m(\omega) - X(\omega)| < \frac{1}{k} \right\}\right) \end{aligned}$$

da cui segue la convergenza q.c. quando il limite all'ultimo membro valga 1. Viceversa se sussiste la convergenza quasi certa, poiché  $\cap_{k=1}^{\infty} A_k \subseteq A_n \forall n$  allora

$$\begin{aligned} P(A) &= P\left(\bigcap_{k=1}^{\infty} \bigcup_{n=1}^{\infty} \bigcap_{m=n}^{\infty} \left\{\omega \in \Omega : |X_m(\omega) - X(\omega)| < \frac{1}{k}\right\}\right) \\ &\leq P\left(\bigcup_{n=1}^{\infty} \bigcap_{m=n}^{\infty} \left\{\omega \in \Omega : |X_m(\omega) - X(\omega)| < \frac{1}{k}\right\}\right) \end{aligned}$$

ma

$$\begin{aligned} &P\left(\bigcup_{n=1}^{\infty} \bigcap_{m=n}^{\infty} \left\{\omega \in \Omega : |X_m(\omega) - X(\omega)| < \frac{1}{k}\right\}\right) \\ &= \lim_{n \rightarrow \infty} P\left(\bigcap_{m=n}^{\infty} \left\{\omega \in \Omega : |X_m(\omega) - X(\omega)| < \frac{1}{k}\right\}\right) \end{aligned}$$

ossia

$$P(A) \leq \lim_{n \rightarrow \infty} P\left(\bigcap_{m=n}^{\infty} \left\{\omega \in \Omega : |X_m(\omega) - X(\omega)| < \frac{1}{k}\right\}\right)$$

da cui l'asserto, essendo  $P(A) = 1$ . □

Essendo

$$P\left(\bigcap_{m=n}^{\infty} \{\omega \in \Omega : |X_m(\omega) - X(\omega)| < \epsilon\}\right) \leq P\left(\{\omega \in \Omega : |X_m(\omega) - X(\omega)| < \epsilon\}\right) \quad \forall m \geq n$$

segue che la convergenza quasi certa implica la convergenza in probabilità.

Quando si abbandona l'ipotesi della somiglianza, occorre fare delle ipotesi più restrittive sui momenti delle v.a.

**Teorema 3.4.** *Siano  $\{X_n\}$  v.a. indipendenti, con media nulla e varianza  $\sigma_n^2 < \infty$ . Se*

$$(3.2) \quad \lim_{n \rightarrow \infty} \frac{\sigma_1^2 + \dots + \sigma_n^2}{n} = 0$$

si ha  $S_n/n \xrightarrow{P} 0$ .

La dimostrazione del teorema è immediata utilizzando la disuguaglianza di Ce-bicev: inoltre si può sostituire all'ipotesi di indipendenza delle v.a. l'ipotesi di incorrelazione. Si può dimostrare che affinché sia vera la (3.2), è sufficiente che

$$\lim_{n \rightarrow \infty} \frac{\sigma_n^2}{n} = 0,$$

ossia che se anche le varianze crescono, al crescere di  $n$ , vadano all'infinito meno velocemente di  $n$ . Allo stesso modo si può dimostrare che affinché sia vera la (3.2) deve accadere

$$\sum_{n=1}^{\infty} \frac{\sigma_n^2}{n^2} < \infty.$$

Questa ultima condizione è anche sufficiente per dimostrare la convergenza quasi certa.